ch5_ Link layer

2hosts بين.Communication مسئولة عن الـ ← Network Layer
Link Layer

Protocol stack

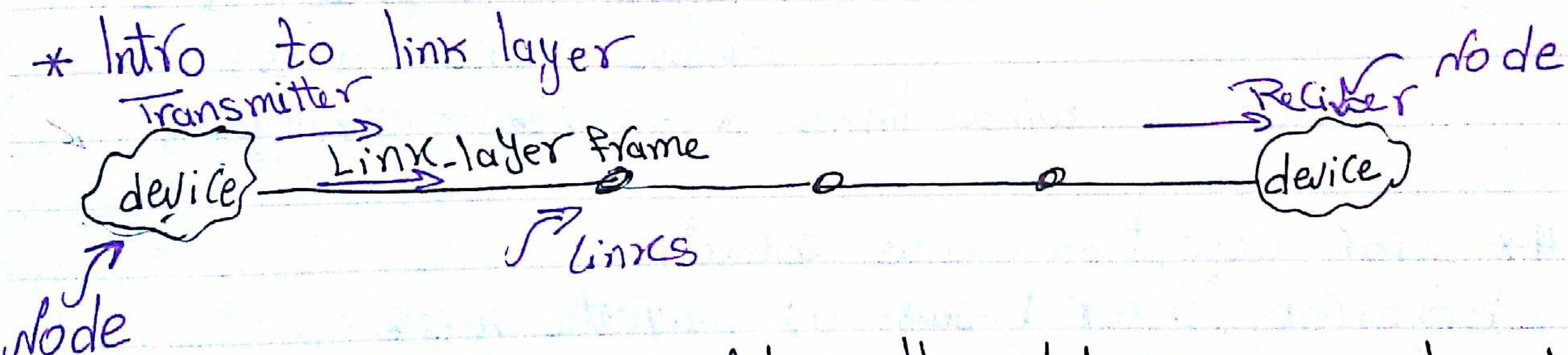
* Types of link_layer channels :-

1_ Broadcast channels

يتوصل بين multiple hosts خلال satellite , wireless LANs, HFC.
وبما أنهم multi hosts فمحتاج medium access protocol لتنظيم الـ transmission
بينهم وعندي حالتين ممكن يكون في central controller للتنظيم أو الـ
hosts بنفسها تنظم الـ trans. بينهم.

2_ Point_to_Point links
Ex: Router to Router _ Pc to Ethernet switch.
وبما أنهم PPP فتنظيم الـ access والـ transmission بيكون بسيط

* Intro to link layer



* Transmitter encapsulates the datagrame in a link layer frame and transmits the frame into the link.

* Services of the link layer:

* الوظيفة الأساسية لها أنها بتنقل الـ datagrame من src.node لـ des. node

1* Framing

مسئولة أنها قبل ما تنقل الـ datagram ع الـ link تعملها
Encapsulation من خلال الـ link-layer frame.

frame:

Data field	header fields

2* Link access

بتحتاج medium access control protocol (MAC) يحدد لقواعد
ويشرف ع عملية نقل الـ frame ع الـ link

✓ Point-to-Point links → MAC's simple or donexist.

✓ Broadcast link → (multiple access problem), MAC protocol's needed to coordinate frame transmission.

3* Reliable delivery

- من خلال acknowledgements & retransmitions
- بنحتاجها مع الـ links اللي فيها معدلات الـ errors عالية
زي الـ wireless links
وبيقل الاحتياج ليها في حالة الـ wired links

4* Error detection and correction

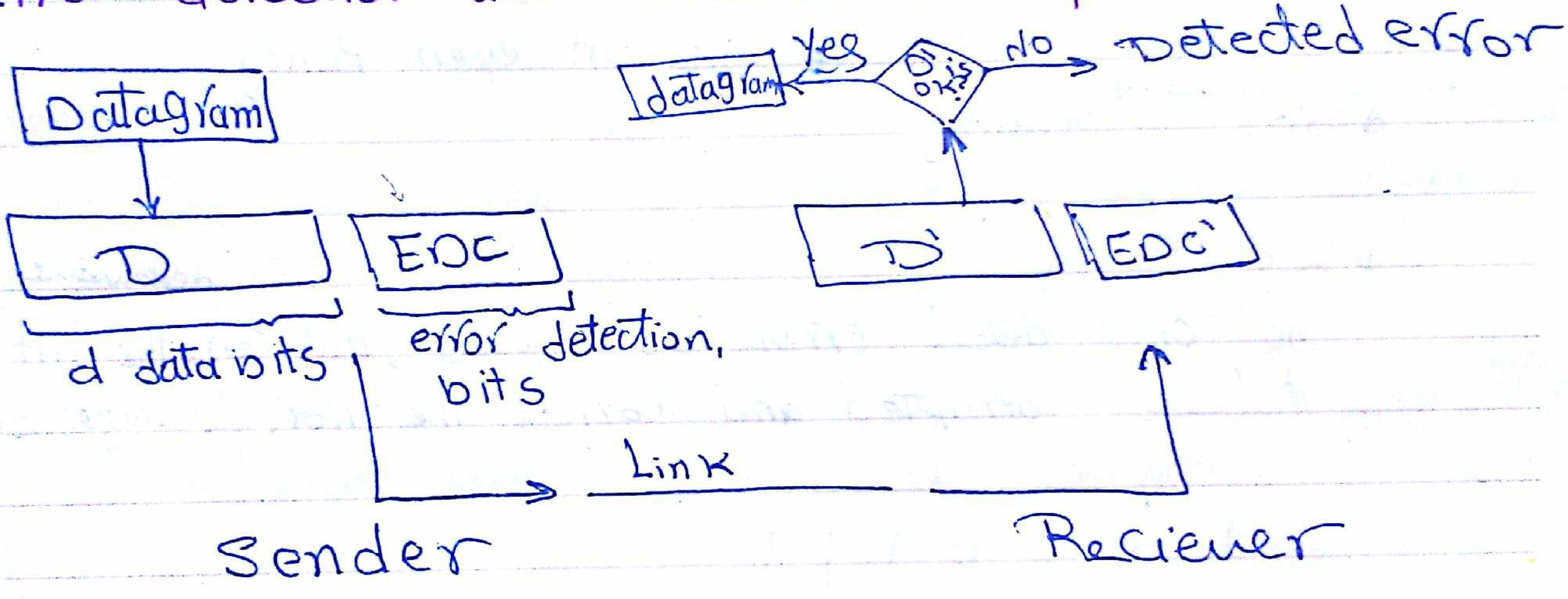
- bit errors may occur because of magnetic noise
- No need to forward datagrame that has error bits so link layer protocols provide mechanism for error detection and correction.
- By, transmitting node include error detection bits in the frame, Recieving node check these bits.
- Error detection in the link layer's implemented By hardware.

* Where's the link layer implemented?
_ Link layer's implemented in a network adapter (Network Interface Card (NIC)).

_ link layer controller : chip, It's part of network adapter, implements link layer services.

_ The link layer's combination of HW and SW.

* Error detection and correction techniques :-



_ Even with the use of error detection bits there still may be undetected bit errors

→ Error detection techniques:
↳ Checksum methods (more used in transport layer)
↳ Cyclic redundancy checks (" " Link Layer)
↳ Parity checks

* Parity checks
  - Single Parity bit for error detection.
    - In even parity schema: No. of 1's in (d+1) bits is even
    - In odd " " " " is odd
* Reciever only count * of 1's in the recieved (d+1) bits
*

* في ال even Parity schema : لو عدد الوحايد فردي ← يبقا ع الأقل في 1bit حصل فيها error أو عدد فردي من ال bits حصل لهم error.
* ولو كان عدد زوجي من ال bits حصل لهم error ⟸ undetected error

$\underbrace{0\ 1\ 1\ 0\ 1\ 1\ 1}_{\text{d bits}}\ \underset{\text{bit parity}}{1}$ → one-bit even Parity

  - 2d Parity
    Not only detect error has occured, but can determine the bit that was corrupted and correct the error. (single bit error)

1	0	1	0	1	1
1	0	1	1	0	0
0		1	1	0	1
0	0	1	0	1	0

Parity error

2d even Parity

* 2d Parity can detect (but not correct) any combination of 2 errors in a packet.

* The ability of the reciever to both detect and correct errors is forward error correction (FEC)

* check Summing

→ at sender

بيعامل الـ segment ع أنها أجزاء 16bit
بيجمع الـ segment bits وبياخد الـ 1'st Complement
وتخزنها في الـ checksum bit

→ at receiver

بيجمع الـ segment bits اللي وصلت.
وبيجمع عليها الـ checksum bit وياخد
الـ 1's Complement لو الناتج طلع أصفار
يكون الداتا وصلت سليمة.

→ why's checksum used at transport layer and cyclic redundancy check used at link layer?

- Transport layer's implemented in SW so it's important to have a simple and fast error detection schema such as checksumming
- Link layer's implemented in HW in adapters, that can perform more Complex CRC operations.

* Cyclic Redundancy check (CRC)

- CRC → Polynomial codes

$\underbrace{\text{D: Data to be sent}}_{}$ — d-bits $\underbrace{\text{R: CRC bits}}_{}$ — r bits (Bit pattern)

r+1 bit pattern → Generator G (Leftmost (significant) bit's 1)

Example
D=1011010010 , d=10 bits
G=110010 , r=5 bits

at Sender

```
G
110010 ) 1101000100
         101101001000000
       ⊕ 110010
         111110
       ⊕ 110010
          110001
        ⊕ 110010
            110000
          ⊕ 110010
               01000  ← R
```

at Reciever

```
G
110010 ) 101101001001000
            :
            :
```

Reminder → if = 000 → No errors happen

*Multiple access links and protocols

* P2P link protocols : PPP _ high level data link control (HDLC)
* broadcast link technologies : Ethernet _ wireless LANs

← لو في أكتر من 2 node حاولوا يبعتوا frames في نفس الوقت ع broadcast link
باخ ال nodes هتستقبل في نفس الوقت وهنا هيحصل collision ع ال frames و.
هتكون lost و BW هيكون wasted وبالتالي محتاجين تدخل
Coordination لل transmission ودة دور ال MAC protocols

→ MAC protocols categories :-
↳ channel partitioning protocols
↳ Random access protocols
↳ Taking turns protocols

* channel partitioning protocols
→ used techniques to partition a broadcast channel's BW :-
① ↳ Time_division multiplex (TDM)
② ↳ Frequency_division multiplexing (FDM)
③ ↳ Code division multiple access (CDMA)

*TDM:

- بيقسم ال time إلى مجموعة Time frames و كل time frame إلى N time slots ، كل time slot بيخصصها assign لواحدة من ال N nodes ، بيراعوا وهما بيختاروا ال slotsize إنها تكون كافية لنقل packet

Time
↓
frames
↓
N-slots

- Adv : TDM eliminates collisions , and fair :-
each node take R/N bps each time frame

- Disadv : * a node is limited to an avg rate of R/N even when it's the only node with packets to send.
* A node must always wait for its turn in trans. seq.

*FDM

- بيقسم ال channel لـ frequencies مختلفة كل واحدة ليها Bw ← R/N وبيعمل assign ال freq. لواحدة من ال N nodes
كأنه بيعمل N channels بمعدل R/N bps

- له نفس مزايا وعيوب ال TDM

*CDMA

- بتعمل assign لـ different codes لكل node ، ال node بتعمل encoding للداتا للرسالة بال code وكل reciever بيستقبل sender's encoded data bits (ال reciever
عارف ~~...~~ ال code ال Sender)

- بتستخدم ف ال military Purpose

* Random access Protocols
- Transmitting node always transmits at the full rate of the channel, R bps. Whe there's a collision, each node retransmits its frame after ~~different~~ random delays

→ used Random access protocols
① → ALOHA Protocols
② → CSMA Protocols
③ → Ethernet Protocol

*Slotted ALOHA
- assume that frame → L bits, time slot size → L/R, transmition → begining of slot, Nodes're synchronized, if frames collide, all nodes detect collision before end of slot.

- ال Node بتبعت ال frame ف بداية ال time slot ، إذا ~~...~~ وصلش collision بيبقى ال frame وصل ونبدأ نجهز ال frame اللي بعده
إذا ~~...~~ حصل detect لـ collision نبدأ نبعت ال frame

مرة ثانية خلال subsequent slot بـ Probability P

- Adv :
  * allows a node to transmit continuously at the full rate R when that node is the only active node.

  * highly decenteralized, as each node detects collisions and independently decides when to retransmit.

  * Slotted ALOHA works well when there's only one active node.

Figure 5.10 → ص 451

Notes on figure :
- successful slot 's in which exactly one node transmits.
- if each node were immediatly retransmit after each collision, → efficiency = 0

- max. efficiency of slotted ALOHA :-

assume :

- كل node تبعث أو تعيد إرسال بعد ال collision بـ Probability P
- N nodes
- احتمالية أن ال slot تكون successful slot هي احتمالية أن node واحدة تعمل transmit والباقي N-1 من ال nodes كلهم fail

$\therefore P_{\text{success slot}} = P^1 * (1-P)^{N-1}$

$\because$ There're N nodes

$P = NP(1-P)^{N-1}$

$\therefore$ efficiency $= NP(1-P)^{N-1}$

$\therefore$ max efficiency $= 1/e = .37$

$\therefore$ effective transmission rate $= .37R$ bps

* ALOHA

- First ALOHA , decentralized , unslotted.

- لما الـ node بيوصلها الـ frame بتعمل transmit فوراً ، لو الـ frame حصلها
collision ← الـ node هتعمل Retransmit فوراً باحتمالية P
خلاف كده ، الـ node بتنتظر الـ frame trans time وبعد كده تبعت الـ frame
باحتمالية P

figure 5-11 → ص 453

- max efficiency of Pure ALOHA :-

$$\therefore P = P(1-P)^{2N-1} = P(1-P)^{N}(1-P)^{N-1}$$

$$\therefore \text{max efficiency} = 1/2e = \tfrac{1}{2} \text{ of slotted ALOHA.}$$

* Carrier sense multiple access (CSMA)

Rules { - Node listens to channel before transmitting → carrier sensing
- a transmitting node listens to the channel while it transmitting if it detects that another node is transmitting an interfering frame, it stops transmitting and waits a random amount of time before repeating the sense and transmit when idle cycle → Collision detection

These rules 're in CSMA and CSMA/CD

→ Why collisions still occur?
figure 5-12

* channel propagation delay

الوقت اللي بتاخده الـ signal عشان يوصلها propagation من node لأخرى
وكل ما الـ delay بيزيد ، فرصة الـ node إنها تعمل carrier sensing
لـ node ثانية بدأت تتصل .

11

* CSMA | CD
* ~~nodes~~ when nodes performs collision detection, it abort their transmission a short time after detecting a collision, to abort them not transmitting a useless frame.

* Operation :- figure 5.13 ص 456
  1_ Prepare datagram from $ NL.
  2_ if the adapter (in node) senses the channel 's idle, it transmits the frame, if do, channel 's busy, it ~~waits~~ waits until it senses no signal.
  3_ While transmitting, the adapter monitors the channel for signal energy coming from other adapter.
  4_ if detect energy, it aborts its transmission
  5_ After aborting, it waits random amount of time and then returns to 2

-> Why we need a random amount of time (Not fixed)?
* عشان لو collision وحصل بينهم في نفس الوقت لو انتظروا نفس المدة هيحصل لهم colliding forever

* Random amount of time -> Back off time

-> what's a good time interval to choose back off time?
* Large time + small # of colliding nodes -> They wait a large time.

* Small time + Large # of colliding nodes -> nodes'll again collide.
* Preferable :- short time <- # colliding nodes 're small
  -> long time <- # colliding nodes 're large

* Binary exponential backoff algorithm
  - used in Ethernet and DOCSIS cable network to solve this problem.

  - when transmitting a frame that has already experienced n collisions, a node choose the value of K at random form {0, 1, 2, ... $2^n - 1$}*
  - in Ethernet, actual amount of back off time = K.512

  - The size of the sets from which K is chosen grows exponentially with # of collisions. so it's called exp. backoff algo.

*Taking turns protocols
→ Examples of used protocols:
1. Polling protocol
2. Token passing protocol

* Polling Protocol

* يتطلب أنه يكون واحدة من ال nodes هي ال master
* ال master بت Poll كل node ع هيئة Round robin fashion.
بمعنى إنها تبعت لل node الاول message بانها تقدر تعمل transmit
ل عدد max frames معين ولما الاولى تخلص تبعت للتانية وهكذا
وال master بيقدر يعرف ال node خلصت ولا لأ بأنه observe
ال channel

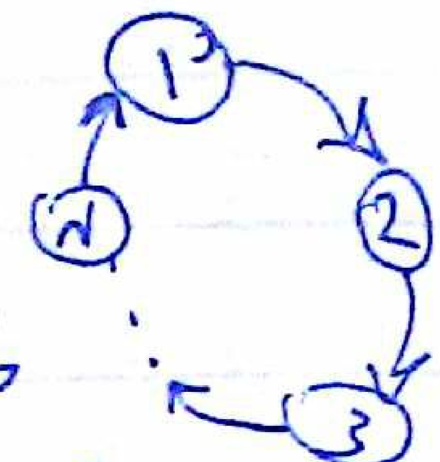
* Adv:
- Eliminate collisions and empty slots
- higher efficiency.

* Dis adv
* Protocol introduces a polling delay (الوقت الي بيعمل فيه polling لل node)
(بيعمل Poll لل nodes inactive)

* If the master node fails, the entire channel becomes inoperative.

* Examples of Polling protocols :-
  ↳ 802.15 Protocol ↳ Bluetooth Protocol

* Token Passing Protocol
  * No master node
  * Special Purpose frame known as a Token is exchanged among the nodes in some fixed order.



* الـ node بتستقبل token بيعمل hold ليها لو
عنده frames بيعملها transmission غير كده بيعمل forward للـ token
لـ node اللي بعده.

→ Adv:
  * Token passing's decenteralized and highly ~~effective~~. efficient.

→ Disadv:
  * The ~~Fail of~~ failure of one node can crash the entire channel.
  * If a node accidently neglects to release the token, then some Recovery procedure must be invoked to get the token back in circulation.

Examples of token Passing Protocol :
↳ Fiber distributed data interface (FDDI)
↳ IEEE 802.5 token ring Protocol

MAC Protocols (summary)

- Channel Partitioning
  - TDM
  - FDM
  - CDMA
- Random access
  - slotted AlloHA
  - AlloHA
  - CSMA
  - CSMA/CD
- Taking turns
  - Polling
  - taken Passing

* Cable access network (case study)
~~fig 5.14 P 461~~
fig P 38 slides

* Connect Cable modems to Cable modem termination system (CMTS) at the cable network headend.

* Data over Cable Service Interface Specification (DOCSIS)
  - specifies cable network architecture and its protocols.
  - uses $\boxed{\text{FDM}}$ to divide upstream & downstream into multiple freq.

  - In downstream: CMTS → all Cable modems.
    There is no multiple access Problem.
    (CMTS → modems)
  - In upstream: (challenging) ‹ Collisions Can occur.
    (modems → CMTS)
  - upstream uses $\boxed{\text{TDM}}$ to divide into intervals of time, each containing a sequence of mini-slots during which modems can transmit to CMTS.

  - CMTS give permission to cable modem to transmit by sending control message (MAP message)

  - How CMTS know which Cable have data ?!
    By having cable modems send mini-slots request frames to the CMTS during a special set of interval ~~time~~ mini slots, that're dedicated for this purpose., They're transmitted in random and collid can occur btw them.

طيب دلوقتي ال Cable modem مش عارف إذا كان ال link هو اللي busy ولا ال frames حصلها Collision

فهو هيفرض إنه لو مستقبلش رد من ال CMTS يبقى أكيد حصل Collision

فحالة إنه حصل Collision: ال Cable modem بيستخدم Back off delay
عشان يعيد إرسال ال frame؟

*Switched Local area networks
→ Link layer addressing and ARP

Switched local area networks

*Link layer addressing and ARP

- Hosts and routers have link layer addresses, network layer addresses as well.

- Address resolution protocol (ARP) provides mechanism to translate IP addresses to link layer addresses.

→ MAC addresses

- Hosts and routers' adapters have link layer addresses (adapters → Network interfaces).

  يعني أن كل interface ل link layer address

- link layer switches' interfaces don't have link layer addresses.

- link layer address is called LAN address, physical address, MAC address.

- MAC address is $\boxed{6}$ bytes long (giving 2^{48} possible address), expressed in $\boxed{\text{hexadecimal}}$ notation.

- MAC address is unique for each adapter. IEEE manages the MAC address space.

- Adapter's MAC address has a flat structure and doesn't change no matter the adapter goes.

- Scenario (sending & recieving frames)

When an adapter wants to send a frame to some destination adapter. It inserts destination MAC address into the frame and then sends the frame into LAN.

A switch broadcasts an incoming frame onto all of its interfaces. So an adapter may recieve a frame that isn't addressed to it.

When an adapter recieves a frame, it'll check to see whether the des. MAC into the frame matches its own MAC address. If it matches, the adapter extracts the encapsulated datagram and passes the datagram up the protocol stack. If it doesn't match the adapter discards the frame.

If a sending adapter doesn't want all the other adapters to recieve the frame, the sending adapter inserts a special MAC broadcast address into the destination address field of the frame.

→ ARP

* what ARP does?

Translate between link layer address and network layer address.

- لو في hostA (src) عاوز يبعت frame لـ host B (des) لمعلومية الـ IP لـ B
فهيضيف للـ frame الـ destination MAC addr هيحصل عليه
عن طريق الـ ARP

IP addr. ——→ [ARP] ——→ MAC addr.

- ARP resolves IP addresses only for hosts and router interfaces on the same subnet.

* How ARP works?

- Each host and router has ARP table.

- ARP table

IP addr.	MAC addr.	T.T.L (when each mapping'll be deleted)

- Scenario → (sending & Recieving at the same subnet)
→ (IP → MAC mapping not found in the table)

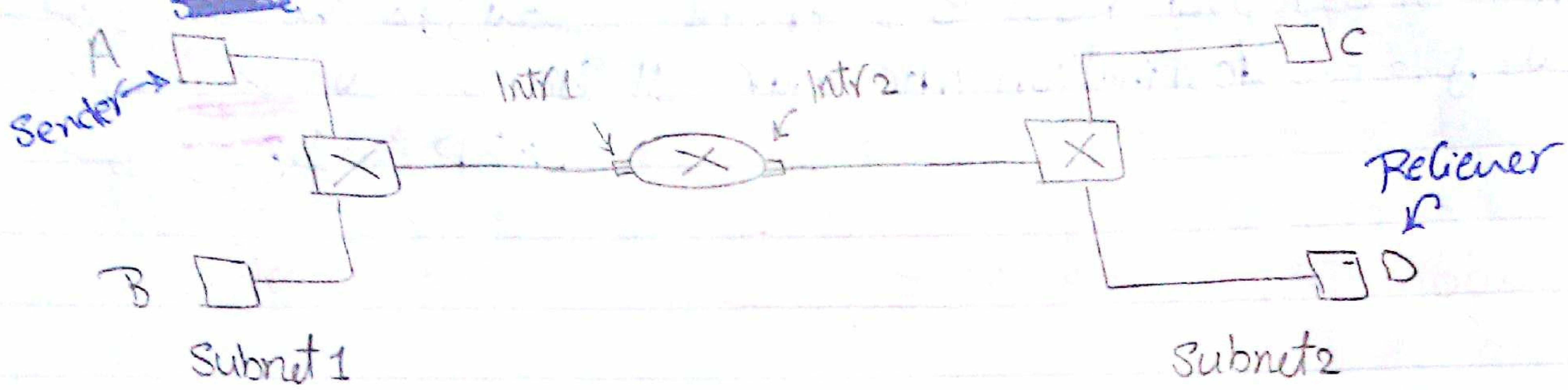
*لو الـ table مفهوش mapping لـ IP معين، الـ sender بيرسل
ARP Packet لكل الـ nodes الموجودة في الـ subnet، وكل node هتقارن
مابين الـ dest. IP addr. اللي في الـ Packet وبين الـ IP بتاعها لو حصل
match بينهم هترد بـ ARP response Packet فيها الـ MAC بتاعها وبالتالي
يقدر الـ sender إنه يحدث الـ ARP table الخاص به.

* Notes: → The query ARP message is sent within a broadcast frame, wheras the response ARP message is sent within a standard frame.

- ARP is Plug and play; ARP table gets built automatically without configuration by admin.

- If a host becomes disconnected from the subnet, it's entry is evenetually deleted from ARP tables.

- Scenario (Sending a datagram off the subnet)



* ال Sender A هيجهز datagram ب dest. IP addr. لل D وهيعملها encapsulation
ب frame بال MAC addr لل intr1 لل router ، وبعدكدة يبدأ
ال switch يعمل broadcast لل frame في subnet1 ، لحد لما ال frame
يوصل ل intr1 هيحصل match ما بين ال MAC بتاعه واللي موجود
في ال frame وبالتالي ال adapter لل intr1 هيفك ال frame
ويوصل لل IP datagram وهنا هيظهر دور ال forwarding table هيعمل
forward لل datagram (اللي ليه IP addr لل D) لل interface المناسب
فال router هيعملها forward ل intr2 ومنه يبعتها
encapsulation في frame جديد بال MAC addr لل D وهيعملها broadcast
في subnet2 .

* Ethernet

→ Ethernet history

1. Ethernet with a hub-based star topology.
2. Switched Ethernet

→ Ethernet frame structure

Preamble	Dest. addr	src addr.	Type	Data	CRC

↳ Data field

* Carries IP datagram (46 لـ 1500 bytes)
* Maximum transmission unit (MTU) of Ethernet is 1500 bytes.

لو الـ IP datagram أقل من 46 byte ← لازم الـ data field تتملى
لحد ما توصل لـ 46 والعملية دي اسمها stuffing

لو الـ IP datagram أكبر من 1500 bytes هيحصل fragmentation

↳ Dest. addr.

* Contains dest. MAC address (6 bytes)

↳ Src. addr.

* Contains src MAC address (6 bytes)

↳ Type

* indicates higher layer protocol type (2 byte)

↳ CRC

* allows the reciever to detect bit errors in the frame (4 byte)

↳ Preamble (8 byte)

* أول 7 bytes ليهم قيمة 10101010 وآخر byte قيمتها 10101011
* وظيفة أول 7 bytes إنهم يعملوا synchronization بين الـ sender والـ reciever
* آخر 2 ~~byte~~ bits في الـ byte الأخيرة (11) بتعمل indication للـ reciever إن الـ bytes اللي بعد كده هتكون data.

_ Notes:

* Ethernet isn't reliable(1)
* Ethernet technologies provide connectionless service to the network layer.(2)

لـ(1) لما reciever B يستلم frame من sender A هيفحصه بـ CRC check ، ولو الـ frame كان فيه اخطاء هيعمله discard من غير مايبعت لـ A أي acknoledgment وبالتالي الـ Ethernet ← not reliable

لـ(2) لما الـ sender A عاوز يبعت datagram لـ reciever B ، A يعمل الـ encapsulation للـ datagram في frame وهيبعته ع الـ LAN من غير أي handshaking في البداية مع B .

*Link layer switches

→ what's the role of the switch?

* To recieve incoming link layer frames and forward them onto outgoing links (forwarding)

_ The switch is transparent to hosts and routers in the subnet.

فمثلا الـ host بيعمل addressing للـ frame بالـ reciever host addr. وليس بعنوان الـ switch intr.

_ The rate at which frames arrive to any of the switches o/P interfaces may exceed the link capacity so switch o/P interfaces have buffers.

→ Forwarding and filtering

* Filtering → بتحدد إذا كان الـ frame هيتعمله drop او forward

* Forwarding → بتحدد الـ interfaces الي هيتوجه ليها الـ frame

* Forwarding and filtering 're done with switch table

→ Switch table

Address	Interface	Time
MAC addr.	Switch intr. that leads to that MA addr.	time at which the entry was placed in the table

→ Scenario (Forward & filter in switch)

نفترض إن frame له dest. addr. معين ووصل عند switch intr x فهنا هتظهر ٣ حالات :-

① أن مفيش entry في ال switch table لل addr. وهنا ال switch هيعمل broadcast لل frame ع كل ال interfaces ماعدا ال x.

② أن فيه entry لل addr. وال interface المقابل له هو ال x وهنا مفيش داعي إن نعمله forward لنفس ال intr. اللي جاي منه فهيعطله filtering وهيعمل drop لل frame

③ أن فيه entry لل addr. وال interface المقابل له هو ال y (≠x) وهنا هيعمل forwarding لل frame لل o/p buffer المقابل لل intr. y

→ How does the switch table get configured in the 1st place?

- Switches 're self-learning; its table's built automatically, dynamically without intervention from a network admin.

- This process (Table configuration)'s accomplished as follows:-

① The table is empty.
② For each incoming frame recieved on an interface, the switch stores in its table (1) MAC addr. of the source, (2) the interface from which, the frame arrived, (3) The current time.
③ The switch deletes an address in the table if no frames 're recieved with that address as src. addr.

- Switches 're plug and play devices as they require no intervention from a network admin.

- Switches 're full_duplex ; any switch interface can send and recieve at the same time.

→ Properties of link-layer switching:-

* Elimination of Collisions ; the switches buffer frames and never transmit more than one frame on a segment at any one time.

* Hetrogeneous links as a switch isolates one link from another, different links on the LAN can operate at different speeds.

* Management
Ex: if an adapter continually sends Ethernet frames, a switch can detect the problem and disconnect the mulfunctioning adapter.

→ Switches VS. Routers

Switches	Routers
→ store and forward Packets using MAC addr.	using network layer addr
→ Layer-2 Packet switch	layer-3 Packet switch
→ Link layer device	Network layer device
→ Plug and play devices	Need Configurations
→ have forwarding table but differ in Computation and Configuration. By self-learning	By Routing algorithms
→ have high filtering and forwarding rates.	have larger Per Packet Processing time
→ To Prevent the cycling of broadcast frames, the active topology ~~of~~ ~~a~~ is restricted to a spanning tree.	Packets don't normally cycle through routers even when the network has redundent Paths. Packets 'r not restricted to a spanning tree and can use the best path between src. and dest.

* Virtual Local area networks (VLANs)

- Switch that supports VLANs allows multiple VLANs to be defined over a single physical local area network infrastructure.

(عاوز أعرف ألترسم VLAN و نفس ال Switch)

- Slide 72 الرسمة

- ال hosts ع نفس ال VLAN ~~بيتعاملوا~~ بيتعاملوا كأنهم ع نفس ال Switch

- Port-based VLAN

* Switch ports grouped so that single physical switch operates as multiple virtual switchs using switch management software (table of Port to VLAN mappings)

* supports traffic isolation:

لو بعتت ال frames من Ports ع VLAN تقدر توصل للـ Ports اللي ع نفس ال VLAN فقط

لو عاوز أعمل Communication مابين VLANs مختلفة :-

* By connecting a VLAN switch port to an external router and configure that Port to belong to ~~these~~ all VLANs.
(vendors ease such configurations by building a single device that contains both a VLAN switch and a router)

- scale up VLANs

* Slide 74 الرسمة

* نفس ال VLAN موجودة ع multiple physical switches

لو هحتاج أعمل Connection مابين ال Switches

* VLAN trunking

↳ approch to interconnect VLAN switches

↳ A special port on each switch is configured as trunk port to interconnect VLANs

↳ The trunk port belongs to all VLANs and frames sent to any VLAN 're forwarded to the other switch over trunk port link.

لي ازاي الـ switch بيعرف الـ frame الموجود ع الـ trunk Port انه يخص VLAN معينة ؟

* IEEE 's defined Ethernet frame format 802.1Q for frames crossing a VLAN trunk.

* 802.1Q VLAN frame format
→ Consists of the standard Ethernet frame with 4 byte VLAN tag added into the header, carries the identity of the VLAN to which the frame belongs.

→ VLAN tag's added by the switch at the sending side of a VLAN trunk.

Preamble	Dest. addr	Src. addr	Tag Protocol id	Tag control Info.	[illegible]	Data	CRC

$\underbrace{\text{Tag Protocol id} \;\; \text{Tag control Info.}}_{\text{VLAN tag (4 bytes)}}$

*Link virtualization : A network as a link layer
→ Multiprotocol label switching (MPLS)

- MPls evolved to improve the forwarding speed of IP routers by using a fixed-length label.
  ↳ allows routers to forward datagrams based on fixed-length label (rather than IP addr.) when possible. (don't inspect IP address)
- The format of a link layer frame that's handled by MPLS-capable router :-

Ethernet header	MPls header	IP header	Reminder of frame

label	EXP	S	TTL
20	3	1	5

EXP: experimental use
S : indicate the end of a series of MPls headers
TTL: time-to-live

- MPls forwarding table distinct from IP forwarding tables.
  ↳ MPls forwarding table (for each router)

in label	out label	dest.	out interface

- MPLS-capable router is called also label-switched router

*How it works?
- label switched routers forwards an MPLS frame by looking up in its forwarding table and then immediatly passing the datagram to the appropriate O/P interface, and need not to extract the IP and perform a look up of the longest prefix match in the forwarding table.

→ MPLS vs. IP Paths

* IP Routing:
- Path to destination determined by destination address alone.

* MPLS routing:
- Path to destination can be based on src and dest address.
- Fast reroute: Precompute routes in case of link failure.

→ slide 82

* Data Center networks

* Thousands of hosts, often closely coupled.
* challanges:
  ↳ multiple apps, serve huge no. of clients.
  ↳ managing / Load balance
  ↳ Networking
  ↳ Rich Interconnection among switches and racks
* Load balancer
  ↳ recieves requests → direct workload to data center → return results to client
  (Hiding data central internal details)